

# Applying multimodal discourse analysis to study image-enabled communication

Jaime Snyder

School of Information Studies,  
Syracuse University  
245 Hinds Hall  
Syracuse, NY 13203

[jasnyd01@syr.edu](mailto:jasnyd01@syr.edu)

## ABSTRACT

A multimodal analytic framework is introduced to contribute to a discourse-oriented study of the creation of visual information. While much visually based research focuses on the image artifact, an ongoing study seeks to shed light on the phenomenon of image creation as a communication practice. This requires a content analytic methodology capable of addressing issues related to modalities of expression and their interaction, co-occurrence and co-deployment during exchange of meaning. Multimodal discourse analysis, an emerging area of discourse studies, is proposed as a valuable contribution to the current study and to the field of information science in general.

## 1. INTRODUCTION

Anyone who has reached for pen and paper during a discussion in order to clarify a thought or prompt a response from a companion has exploited the potential of image-making to enhance communication. Because images and visual information enable exchange of meaning across a range of contexts, they play an increasingly important role in how we work and communicate, in both face-to-face and virtual environments.

Marks on a napkin or sketches on a white board are information artifacts embodying a particular type of image-enabled communicative practice that plays a specific role in the exchange of meaning between individuals. These spontaneously created visualizations can anchor, bridge, and facilitate the flow of information at crucial moments in a conversation. *Image-enabled discourse* is the term used to refer to this phenomenon in a broad sense, and *ad hoc visualizations* (i.e. napkin drawings) are one type of image-enabled practice.

In the context of conversations, image creation is an interactive process that generally results in the creation of an artifact, but also includes the motivation or need for the image, the deployment of the image in a specific context, and the reception of the image within an overarching communicative structure [1]. When we focus predominately on the content of that image (either through automated analysis or more qualitative interpretation), we

run the risk of generating static analyses of graphical content in which the image is seen as a *fait accompli*, rather than evidence of an interactive process of communication. By improving our understanding of the dynamics of image-enabled conversations, we can build better tools to help people engage in more effective communication

The purpose of this poster is to introduce two approaches to the analysis of multimodal communication that can provide the necessary framework for an investigation into the creation of visual information during face-to-face conversation. Norris' approach to multimodal interactional analysis and Baldry and Thibault's techniques for multimodal text transcription both provide guidance for the operationalization of analytic methods necessary for a discourse-oriented study of image-enabled communication.

## 2. RESEARCH QUESTIONS

Recent research focused on the creation of images within the context of face-to-face conversation seeks to address the following research questions:

- RQ1: In what ways does the ad hoc creation of images contribute to communication during face-to-face conversations?
- RQ2: What communication strategies are associated with the creation and use of ad hoc visualizations in these interactions?
- RQ3: Which characteristics of visual modes of communication are most salient to the use of ad hoc visualization strategies?

Addressing these questions requires a perspective that recognizes the constructed and dynamic nature of the exchange of meaning between individuals across a variety of modalities. Multimodal discourse analysis, an emerging methodology in the field of discourse linguistics, is introduced to provide an analytic basis for this investigation.

## 3. BACKGROUND

Although no branch of discourse studies specifically invites extending theories into the realm of image-enabled communication, an emerging subfield of discourse linguistics is highly relevant for this investigation. Multimodal discourse analysis is the study of the intersection and interdependence of various modalities of communication within a given context. Researchers in this area seek to identify the influence of mode on

meaning within a given context, focusing on co-occurrence and interaction between multiple semiotic systems [2].

Generally speaking, *mode* refers to a distinct semiotic system for expressing meaning using specific conventions [3,4]. At the heart of most work in the area of multimodal discourse is the principle that communication occurs across more than a single mode and is therefore inherently *multimodal* [4.5].

According to social semiotic theories related to modality, the form or format of an expression plays a specific role in the communicative power of a sign [4.5]. Modality cues are used in the creation of meaning, referred to as “motivated signs” in the vocabulary of social semiotics [5]. And “...any semiotic mode, even smell, can be conceived of as a loose collection of individual signs, a kind of lexicon, or a stratified system of rules that allow a limited number of elements to generate an infinite number of messages” [6] although the precise nature of those rules or conventions of use can be more or less specific [4].

The generalized definition of modality as a unique semiotic system allows us to go beyond superficial distinctions (such as media or format) to a more complex understanding of how a mode of expression affects the contextualized exchange of meaning(s). Norris highlights this, stating “different communicative modes possess different materiality” [7]. She attributes audible materiality to spoken language, despite its being neither visible nor enduring, while gesture is visible, but also quite fleeting. Print is visible and enduring, as is physical layout of objects. In the context of image-enabled discourse, there are properties of ad hoc visual communication, specifically mark-making, that are uniquely suited to providing the requisite conditions for a person to employ certain communicative practices. For example, drawing naturally has the affordances of being persistent, tangible and visible.

## 4. ANALYTIC FRAMEWORKS

Two specific approaches to analyzing multimodal interactions can be applied in an analysis of image-enabled conversations to identify salient aspects of communication and the role that modality plays in the exchange of meaning.

### 4.1 Multimodal interactional analysis

In her approach to the analysis of the interaction of multiple modes of communication is a single context, Norris focuses on “what individuals express and react to in specific situations, in which the ongoing interaction is always co-constructed” [7] She is not just looking at verbal expressions, but at other types of indicators such as head position, body position and layout of objects or spaces to reveal the ways in which this co-constructed is built. Her analysis of multimodal interactions is based on the concepts of *awareness* and *attention*. She clearly states that “Awareness/attention comes in degrees, and a person may be phenomenally aware of something without paying much attention to it.”

One of the key concepts offered by Norris that can help clarify the unique communicative qualities of ad hoc visualizations and begin to help us understand how this mode interacts with other types of expressions is the distinction between *embodied* and *disembodied* modes of communication. Music, for example, can be either embodied or disembodied. If a radio is playing in the background while a couple is sitting at a table having breakfast, music is seen as a disembodied mode of communication, not receiving a lot of

attention, but the couple will probably have some awareness of it. However, if one of the participants in the conversation breaks into song, music becomes an embodied form of communication, bringing a different level and degree of attention and awareness, not unrelated to the fact that this could be seen as an unusual or unconventional occurrence.

### 4.2 Multimodal transcription

Baldry and Thibault’s approach to multimodal text transcription follows Halliday’s definition of text as “living language” regardless of whether it is spoken, written or takes another medium of expression [8]. They recognize that “Different semiotic modalities make different meanings in different ways according to the different media of expression they use.”

Multimodal transcription is specifically designed to retain relationships between unique modes of expression in order to retain evidence of differences as well as things like co-deployment. It allows *phasal* expressions (i.e. time- or series-based expressions such as gesture) to be recorded and transcribed alongside *clustered* expressions (i.e. groupings or sets of static expressions such images in a magazine spread or elements on a web page).

Additionally, Baldry and Thibault’s techniques highlight the multi-level aspects of meaning making across modalities. They use the concepts of *context of situation* and *context of culture* to delineate important distinctions between modalities. They also examine relationships between individual multimodal *texts* and multimodal *genres*. And like Norris, they acknowledge that *primary* and *secondary* genres exist within any multimodal text.

## 5. CONCLUSION

Within the field of information science, there is need for more robust methodology that addresses the role that information creation (and visual information, in particular) plays in the exchange of meaning during interactions between individuals. There is great value in having the ability to differentiate images not just by form or format but also by role in communication.

The methodological approaches presented here, multimodal interaction analysis and techniques for multimodal transcription, provide a practical basis for investigating the role drawing plays in conversation. Originating from a discourse perspective, these approaches capture the interaction and co-deployment of multiple modes of expression, and allow researchers to track the influence of these differences in exchange of meaning. Use of these methodologies will help to increase understanding of image-enabled communication and allow us to better exploit this innate human communication practice when building tools and systems.

## 6. REFERENCES

- [1] Ware, C., Information Visualization: Perception for Design. 2000, San Francisco: Morgan Kaufman.
- [2] Royce, T.D. and W.L. Bowcher, eds. New Directions in the Analysis of Multimodal Discourse. 2007, Lawrence Erlbaum Associates, Publishers: Mahwah, NJ.
- [3] Bateman, J.A., Multimodality and Genre: A Foundation for the Systematic Analysis of Multimodal Documents. 2008, New York, NY: Palgrave Macmillan.

- [4] Kress, G. and T.V. Leeuwen, *Multimodal Discourse: The Modes and Media of Contemporary Communication*. 2001: Arnold.
- [5] Kress, G. and T.V. Leeuwen, *Reading Images: The Grammar of Visual Design*. 2006: Routledge.
- [6] Levine, P. and R. Scollon, *Discourse and Technology: Multimodal Discourse Analysis*. 2004, Washington, D.C.: Georgetown University Press.
- [7] Norris, S., *Analyzing multimodal interaction: A methodological framework*. 2004, New York and London: Routledge.
- [8] Baldry, A. and P.J. Thibault, *Multimodal transcription and text analysis*. 2006: Equinox.